Creation of *Vibrio cholerae* strains to test the induction of the Feo iron transport system using the recombination-based in vivo expression technology (RIVET) vector

Suzanna Maupin

University of Texas at Austin

Supervising Professor: Shelley Payne, Ph.D.

_____

Signature Date

Table of Contents

Abstract:

Cholera is a diarrheal disease caused by the bacterium *Vibrio cholerae.* Iron is important for survival for this pathogen, both as a regulatory mechanism and as part of enzymes. Although the iron-transport mechanisms of *V. cholerae* have been studied in great detail, it is still unclear when these mechanisms are active during infection. But, there is now a method to test the expression of these mechanisms. This method is known as the recombination-based in vivo expression technology (RIVET) vector. This is a new technology with which one can look at the induction of transcription of different genes during infection of infant mice. I have inserted the *feo* gene promoter, which is a ferrous iron uptake mechanism, into this vector. I have then incorporated this vector into two cholera strains, so that we can better understand when this particular iron uptake mechanism is expressed in vivo. These strains will be able to successfully identify when the iron transport genes are expressed during infection.

Background:

Basic information about *V. cholerae*

Cholera, a severe diarrheal disease caused by the bacterium *Vibrio cholerae*, is an important disease both historically and today, especially in developing countries. It severely affects developing countries where clean water is not readily available (Sack et al., 2006). It can lead to a debilitating enteric disease, and treating it can be difficult both financially and physically. One of the tactics that researchers are using to attempt to better understand and perhaps fight this pathogen is to look into the iron binding and transport mechanisms of these bacteria. Since iron is important for the survival of *V. cholerae*, understanding the way that these iron uptake mechanisms work could be incredibly beneficial to preventing the spread and reducing the pathogenicity of this disease.

V. cholerae is an enteric bacterial pathogen within by the *Vibrio* genus. *Vibrio* are gram-negative facultatively anaerobic fermentative rods. The pathogen that we are focusing on out of this large group of ten species is one species, *V. cholerae*. This species is identified by the O antigen of its lipopolysaccharide (LPS), and there are over 140 serogroups of *V. cholerae* that are characterized in this way. The only known serotypes that cause disease are the ones that also produce the cholera toxin, serotypes O1 and O139 (Murray 2005). The cholera A-B toxin is encoded on the bacteriophage CTX Φ. Five B subunits form a ring and bind to ganglioside receptors on human enteric epithelial cells. The A subunit then reacts with G proteins controlling adenylate cyclase, leading to an increase of cyclic AMP in the epithelial cell. The increased cAMP leads to a loss of water and ions from the cells, which is the reason that patients who have contracted this disease have very watery diarrhea. *V. cholerae* uses pili that attach to these

epithelial cells, as well as chemotaxis proteins, to prevent being washed out with the enormous volume of water that is being passed through the intestine (Faruque et al., 1998).

Through this infectious mechanism, the patient will have symptoms of watery diarrhea and vomiting. Because of the loss of water, the patient can also have dehydration, and can go into shock without fluid replacement. One of the symptoms that can help to diagnose cholera from other enteric diseases is the passage of odorless stools, since the majority of the stool is made up of water being lost from cells. It is quite a common disease in developing countries with poor hygiene, as it is passed through fecal-oral transmission. The treatment is fairly simple; replacement of fluids and electrolytes and the administration of antibiotics such as tetracylcine, and with correct treatment, the death rate goes down from 60% to 1% (Murray 2005). It has been shown that *V. cholerae* O1 strains can bind to chitin, which explains the close association of the disease with eating raw shellfish (Reidl and Klose 2002). The infectious dose of bacteria is 10^8 organisms, since the normal strains of *V. cholerae* are not particularly acid-tolerant and therefore need a large infectious dose to be able to survive passage through the stomach's low pH. This is why contaminated food and water is usually the vector for passage of the disease, since it would be difficult for person-to-person transmission to result in such a large infectious dose. There are vaccines being developed against this bacterium, and the current vaccines lose their effectiveness over a period of only 3 years (Murray 2005). Therefore, it is important to research into other areas of protection from this important disease, such as the possibility of stopping the disease through knowledge gained from their iron transport mechanisms.

Iron transport in *Vibrio cholerae*

The bacteria *V. cholera*e has reduced virulence when the amount of iron available is reduced (Payne and Finkelstein 1978). This demonstrates how important iron transport is to this

pathogen, and it also shows that being able to understand iron transport mechanisms could give incredibly valuable insight into methods of reducing the number of people who have to suffer from enteric diseases. Many studies have been done showing that iron is incredibly important to the ability of these bacteria to infect cells, such as by increasing the amount of iron in animal cells, and showing the virulence is increased (Payne and Finkelstein 1978). Yet, even now, we have a fairly limited understanding of all of the iron uptake mechanisms and at what point they are used.

It is not easy for a bacterial pathogen to uptake iron while in the human body. Although there are usually around 5 grams of iron in the human body, about 75% of it is stored in cells as hemoglobin. The majority of the iron not stored as hemoglobin is stored as ferritin in the liver, with a small remainder being complexed to iron-binding proteins such as transferrin or lactoferrin (Neilands 1995). Survival within a host can be very difficult for a pathogen which needs iron. Because these bacteria do need iron, they have developed a large array of iron transport mechanisms. One way that bacteria are able to retrieve iron from the host environment is through the use of siderophores. These are iron chelating compounds that are used to bind to iron in the ferric (+3) state. Ferric iron is found in water with high levels of oxygen, such as the human body, and it is insoluble in water. Siderophores bind to ferric iron to form soluble Fe^{3+} complexes, and the microbe is then able to bind to the siderophore and take the iron into the cell (Neilands 1995). *V. cholerae* is a facultative anaerobe, so while in an aerobic environment it uses iron to help with cellular respiration to make ATP. *V. cholerae* can also survive anaerobically, and even in this environment there are many reasons that this microorganism needs iron, such as the reduction of ribotide precursors of RNA and the production of haem. Also, bacteria use iron

in various enzymes, and it is an important regulatory co-factor for bacterial cells. There are many reasons that bacteria need iron that are crucial to successful survival.

Iron uptake mechanisms

V. cholerae has a variety of ways that it is able to acquire iron from the external environment. It is able to make a catechol vibriobactin, a siderophore, which is made up of a norspermidine backbone with three dihydroxybenzoate moities. This is the only siderophore that has been described for *V. cholerae*. Besides making its own siderophore, this bacterium can also take up siderophores made by other species, such as enterobactin made by *Shigella* and *E. coli*, as well as fluvibactin and agrobactin that have structures similar to vibriobactin (Andrus et al., 1983).

V. cholerae has a poorly understood ferrous iron transporter called Feo. Although much is known about the homologous Feo transporter in *E. coli*, very little is known about the one in *V. cholerae.* In both *E. coli* and *V. cholerae*, the transporter is made up three proteins, FeoA, FeoB and FeoC. FeoA and FeoB in *V. cholerae* have 40% homology with the *E. coli* proteins of the same name, while FeoC is the same length in *V. cholerae* as in *E. coli* but only has 11% amino acid homology (Wyckoff et al., 2006). FeoB is a pore protein, which has a hydrophobic region that passes the membrane 7 times. FeoA is a small hydrophilic protein that is associated with FeoB in some unknown way to help FeoB perform its function. FeoC is another small hydrophilic protein of unknown function. The *feoA* and *feoB* genes are adjacent to one another and found within an operon, showing that FeoA function may be associated with FeoB. Mutations in *feoA* do not reduce iron-uptake activity as drastically as do mutations in *feoB*, demonstrating that the FeoB protein is the major part of this transport mechanism (Cartron et al., 2006). Since *feoA* shows some homology to the Src homology 3 domain, or the SH3 domain, it

may help FeoB perform its function by acting to help the G-protein on FeoB by activating its GTPase (Cartron et al., 2006). The SH3 like region on the FeoA protein interacts with a G-protein on FeoB, which regulates the uptake of iron through the FeoB transmembrane pore (Cartron et al., 2006). Through tests done with radioactive iron, it was found that the transport of iron by Feo was stimulated by addition of the reducing agent ascorbate, which is consistent with the hypothesis that Feo transports ferrous iron (Wyckoff et al., 2006).

Another iron transporter of *V. cholerae* is FpbABC, which consists of a periplasmic binding protein and a cytoplasmic protein of the ABC transporter family. This transport system, made up of three proteins, is closely related to one in *M. haemolytica*, and is an example of one of the basic active transport systems within cells. The transport of iron by this mechanism was inhibited by the addition of ascorbate, consistent with the hypothesis that it is a ferric transporter (Wyckoff et al., 2006).

V. cholerae is also able to uptake and use haem. Researchers have identified the *h*aemin *ut*ilization, or Hut, proteins. HutA is a 77 kDa outer membrane protein which acts as a receptor for haem and HutR is very similar to HutA. When a mutant was formed that had *hutA* and *hutR* genes that did not function, the bacteria could not use hemoglobin as an iron source and could only weakly use haem (Mey and Payne 2001). HutB is a 26 kDa inner membrane protein, which is potentially a periplasmic binding protein (Henderson and Payne 1994). HutC and HutD make up the ABC transporter that transports the haem across the cytoplasmic membrane (Selinger 2001). The HutR transport system for haem presumably works in much the same fashion. It was found that both single mutants in vibriobactin uptake and mutants in haem uptake had little effect on the ability of *V. cholerae* to colonize the intestines, which shows that there are a diversity of iron transport mechanisms used by these bacteria (Henderson and Payne 1994).

TonB dependent transport systems are also important in *V. cholerae*. TonB-ExbBD complexes transduce energy to the outer membranes of Gram-negative bacteria (Postle, 1990). TonB is anchored in the membrane by a single membrane spanning domain. Mutations of *tonB* lead to Gram-negative bacteria that are unable to transport vitamin $B_{12,}$ iron siderophores, or haem. The two TonB systems are associated with vibriobactin, enterobactin, ferrichrome and haem transport (Occhino et al., 1998).

V. cholerae iron transport is also regulated by Fur. When there is a high level of iron in a cell, iron bound to Fur will bind to promoters of iron-regulated genes, preventing their expression, while without iron, these Fur proteins do not bind to and suppress the expression of Fur-regulated genes. Fur mutants, which are unable to correctly control their iron uptake mechanisms, demonstrate a variety of phenotypes. It has been found that Fur, in addition to controlling iron uptake mechanisms, also controls the expression of *rhyB*. This inhibition of *rhyB* leads to the expression of genes for motility and chemotaxis. Other proteins that are repressed by Fur are IrgA, an enterobactin receptor, and IrgB, which is required for transcription of *irgA*, (Watnick et al., 1998).

As can be seen, there are many different mechanisms for iron uptake and transport in *V. cholerae*. It may seem that the rationale for exploring all of these different mechanisms is for a purely scientific standpoint and that the knowledge gained could not be used for medical purposes. Yet, if you look at all of the virulence factors that are regulated by iron, it is easy to see that these proteins would be good targets to produce vaccines, since they are highly expressed during infection with the pathogen. Therefore, besides the pure knowledge from studying these iron transport mechanisms, there is also a practical medical aspect. There is much more work to be done on iron transport systems in this bacterial pathogen. There is a need to find the way that

the Feo transporter works, and which iron transport systems actually perform during infection of mammalian cells. The project that I will be working on is related to these two issues.

The RIVET vector

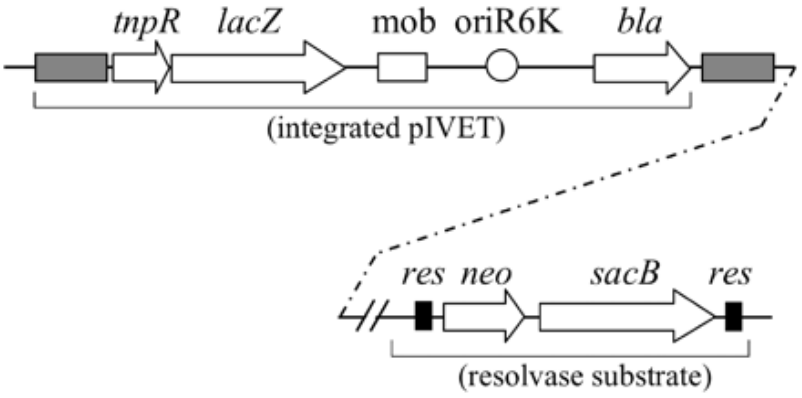

Figure 1: The RIVET vector
This is a diagram of the integration of pGOA1193 (the integrated pIVET) into the genome with the pRES resolvase substrate. Genes are shown as arrows and labeled (*tnpR*: resolvase substrate, *lacZ:* β galactosidase, *bla*: ampicillin resistance, *neo*: kanamycin resistance, *sacB*: sugar sensitivity). The *res* sequences are the filled in rectangles, and the mobilization (mob) and origin of replication (oriR6K) are the open rectangle and closed circuit. The grey filled rectangle is the fragment cloned into the pIVET derivative allowing for homologous recombination. (Figure from Osorio et al., 2004)

In 2004, Osorio et al. improved a way to screen *V. cholerae*'s expression of genes. This method is called recombination-based in vivo expression technology, or RIVET. This builds off a previous method called IVET, or in vivo expression technology. In this new method, two cassettes, pGOA1193 (the integrated pIVET in Figure 1) and pRES (the resolvase substrate) are inserted into the bacterial DNA. In this method: first, the resolvase substrate, such as pRES or pRES1, is conjugated into the *V. cholerae* strain of interest through mating with *E. coli* bacteria containing the resolvase plasmid. After this mating has occurred, selection is performed on the colonies for bacteria that have successfully incorporated the entire vector into their genome through a double crossover. The selection occurs through testing for kanamycin resistance, since the gene for this resistance, *neo*, is only found on the resolvase.

The DNA of interest, in my case the *feo* promoter, must also be cloned into the pGOA1193 vector. This vector contains the resolvase, *tnpR*, as well as genes which confer the ability to use lactose, *lacZ*, and a gene conferring resistance to ampicillin, *bla* (Figure 1). This vector containing the gene of interest is then conjugated into the *V. cholerae* strain containing the

resolvase created earlier. Once this strain of *V. cholerae* with the vector incorporated is formed, it can be used to test whether the gene of interest is expressed in vivo (Osorio et al., 2005). The *neo* and *sacB* genes found in the resolvase substrate mean that we can test whether the gene of interest is expressed. As stated earlier, *neo* leads to kanamycin resistance, and *sacB* leads to sucrose sensitivity. This is because when SacB produces toxic products in the presence of sucrose. Therefore, if the bacteria have this cassette incorporated into their genomes, they will be resistant to kanamycin and sensitive to 10% sucrose. If the gene of interest is expressed, however, a different pattern of resistance will be seen. When the gene of interest is expressed, the resolvase TnpR will be produced and will lead to recombination between the *res* sequences which flank the *neo* and *sacB* genes. This excises these genes, leading to bacteria that are kanamycin sensitive and sucrose resistant. After obtaining a strain that has successfully incorporated this cassette, there can be many tests done to ascertain when the gene of interest, *feo*, is expressed. The resolution of the cassette is essentially irreversible, so after the gene of interest is transcribed, the bacteria will be permanently kanamycin sensitive and sucrose resistant. Therefore, the bacteria can be used to infect mouse models, and the resistance pattern of the bacteria will show whether the gene was transcribed during the infection.

This method has been used for large-scale screening for genes in *V. cholerae* by Osorio et al. They were able to successfully use this technique to find genes that were turned on in *V. cholerae* only after the bacteria infected mice. After they made a library of strains that had the RIVET vector inserted into different genes, they screened all of the strains so that they could be sure to look at genes that did not resolve during growth in L-broth. They then inoculated their strains into mice and after a 24 hour infection plated serial dilutions on sucrose plates. Afterwards, they took the sucrose resistant strains and found the genomic insert into the plasmid

through plasmid DNA purification and electroporation of the plasmid into *E. coli*. The *V. cholerae* DNA found in the plasmid was then sequenced. In this way they were able to obtain forty fusion strains that expressed particular genes to a greater extent during infection than in vitro. These genes include three genes related to chemotaxis, thirteen metabolic genes, and twelve genes involved in transport across the envelope (Osorio et al., 2005).

Having a strain such as this one could be an invaluable tool for obtaining important knowledge about the Feo transport system. The strains that contain the Feo promoter can be used to discover when this iron-transport mechanism is active, such as when the bacteria is grown in vivo or in vitro. Also, the *V. cholerae* strains containing the resolvase substrates can also be used in the analysis of other genes with the RIVET vector, since these strains can easily be conjugated with pGOA1193 vectors containing other cloned genomic fragments. These strains will be used to determine under what circumstances different *V. cholerae* genes are used.

Materials and Methods:

Bacterial Strains and plasmids

The *E. coli* SM10 λ-pir and DH5α λ-pir competent cells made by Elizabeth Wyckoff carrying the appropriate plasmids were mated with the various *V. cholerae* strains to result in the final strains of interest. The *V. cholerae* strains used, also obtained from Elizabeth Wyckoff, were ALV101 and O395. The plasmid pWKS30 was used to transfer the *feo* DNA into pGOA1193. pRES, pRES1 and pGOA1193 were plasmids received from Andrew Camilli at Tufts. These were integral to producing the final RIVET strains. The sequences for pRES and pGOA1193 are in appendices 1 and 2.

Primers

Feo.908.Bgl 5´ TAAAGATCTCTCTTATTTCAGTCAGCGTTTC 3´

Feo.1852.rev.Bgl 5´ GGGAGATCTCCCATCACCATCAGTTTCTTTC 3´

Vc.lac.147 5´ CCAGCGGTTATGAGGTGTATGG 3´

Vc.lac.4321.rev 5´ GCGGAGCGTGTGTGATTTTAGG 3´

RIVET.4018.rev 5´ AGCGAAAGGAAACAATGTCGTG 3´

RIVET.6123 5´ GATTGTCGCACCTGATTGCC 3´

RIVET.9276.rev 5´ TCAAAAACGCCGAGTCCATAAC 3´

Media, reagents and growth conditions

Bacterial strains were stored at -80° C in tryptic soy broth (TSB) and 20% glycerol. All *E. coli* and *V. cholerae* strains were grown in Luria broth (L broth) or on Luria agar (L agar) at 37° C. Antibiotics were used at the following final concentrations: 100 µg ampicillin per ml, 10 µg streptomycin per ml and 50 µg kanamycin per ml.

Plasmid DNA Isolation

Plasmid DNA was isolated using the Sigma plasmid minipreparation procedure, using buffers provided in the kit and TE´ (10 µM Tris Cl and 0.1 µM EDTA). First, 3 ml of culture was centrifuged, the supernatant removed, and then 200 µl of resuspension buffer was mixed with the pellet by pipetting up and down. Then, 200 µl of lysis buffer and 350 µl of neutralization buffer were added to the cells, and both were mixed through inversion of the tube. This mixture was centrifuged for ten minutes at 13200 rpm, and the supernatant was added to the Sigma minipreparation columns, centrifuged for 1 minute at 13200 rpm, and the flow through was discarded. The DNA left on the column was washed with 500 µl of optimal wash buffer, then centrifuged one minute at 13200 rpm and the supernatant discarded. Then the column was washed with 600 µl of wash buffer, then centrifuged one minute at 13200 rpm and the supernatant discarded, and then the column was centrifuged one minute at 13200 to dry. The DNA was eluted off the column into a clean Eppendorf tube with 100 µl of TE´, which was left on the column for one minute and then centrifuged another minute at 13200 rpm.

Transformation protocol

Plasmid DNA was transformed into *E. coli* strains cells using the heat shock method. First, the competent cells frozen at -80° C were put on ice until thawed. Then, 2 µl of the plasmid, obtained through the plasmid DNA isolation described earlier, was added to competent cells, such as SM10 λ-pir and DH5α λ-pir, and this mixture was left on ice for one hour. After this time, the cells were subjected to a one minute heat shock at 42° C so that they would uptake the plasmid, and then were put in 1 ml L broth to shake at 37° for one hour. After this step, the cells were plated on LB agar plates containing the correct antibiotics for selection. Three plates for

each transformation were used that received either 10, 100 or 1000 µl, so that adequate separation of colonies could occur.

Polymerase Chain Reaction (PCR) methods

Pfx PCR: a PCR using the Pfx enzyme was done to amplify the *feo* promoter from Lou15 DNA. This was done using the Platinum Pfx polymerase from Invitrogen, and the associated buffer. The Pfx PCR was done in 50 µl. The protocol is as follows: 2 µl of *V. cholerae* Lou 15 DNA, 10 µl of Pfx 5X buffer (total dilution of 1X), 7.5 µl of 2 mM dXTP (0.3 mM final concentration of each), 0.75 µl of the 20 µM forward primer (Feo.908.Bgl) and 0.75 µl of the 20 µM reverse primer (Feo.1852.rev.Bgl) (0.3 µM final concentration each), 1 µl 50 mM $MgSO_4$ (1 mM final concentration), 27 µl sterile water, and 1 µl of Pfx DNA polymerase (2.5 units final concentration) were used to run the PCR, for a 50 µl final volume. The PCR was run at 94 degrees for two minutes as a preheat, and then was run for 30 cycles of 94 degrees for 15 seconds, 55.8 degrees for 30 seconds, and 68 degrees for 1 minute 15 seconds. After the thirty cycles, the reaction was completed by another elongation step of 68° C for five minutes.

Taq PCR: the PCR using Taq was used to test for the desired products of the mating and transformation. Templates were 100 µl of an overnight culture centrifuged and resuspended in 1 ml water. Taq and the associated buffers were obtained from New England Biolabs. It was run as follows: 1 µl of template, 1 µl of the forward and reverse primers both at 20 µM (0.2 µM final concentration), 2 µl 2 mM dXTPs (0.2 mM final concentration), 2 µl 10X buffer (1 X final concentration), 0.2 µl Taq polymerase and 12.8 µl of water (to bring the total volume up to 20 µl). The PCR was run at 94 degrees for two minutes as a preheat, and then was run for thirty cycles of 94 degrees for 45 seconds, optimum annealing temperature (usually 57.8) degrees for

45 seconds, and 72 degrees for 1 minute per 1000 base pairs. After the thirty cycles, the reaction was completed by another elongation step of 72° C for five minutes.

Gel extraction

The gel extraction was also done using a Sigma kit. The band containing the DNA cut from the gel was solubilized with three gel volumes of the gel solubilization solution, heated at 50° C for five minutes, and vortexed. The binding column was prepared with 500 µl of the column preparation solution and centrifuged for one minute. Then, one gel volume of isopropanol was added to the solubilized gel fragment, and the solution is added to the binding column and centrifuged at 13200 rpm for one minute, 700 µl at a time. Finally, the column was washed with 700 µl of the wash solution, centrifuged for one minute at 13200 rpm, centrifuged for one minute at 13200 rpm to dry, and the DNA was eluted off the column with 50 µl of the elute solution.

Ligation

Ligation was done using T4 ligase from New England Biolabs. The reaction was carried out in 10 µl with 1 µl 10X NEB4 ligation buffer (final concentration 1X) and 1 µl T4 DNA ligase. For the ligation of the *feo* promoter into pWKS30, ATP was needed, so the final three components were 2 µl of the vector (pWKS30 SmaI fragment), 5 µl of the *V. cholerae feo* promoter and 1 µl of 5 mM ATP. The ligation of the *feo* promoter into the pGOA1193 vector instead used 1.5 µl of vector and 6.5 µl of the *feo* promoter insert, since ATP was not necessary. Both were incubated at 16° C overnight.

Restriction enzyme digest

All restriction enzyme digests were done using New England BioLabs enzymes and buffers as indicated in Table 1.

In a 10 µl reaction: 1 µl 10X specific NEB buffer (1X final concentration) 1 µl enzyme (10-20 units), 6-8 µl of DNA, water was added to a final volume of 10 µl.

Table 1. Restriction enzymes and buffers:

Restriction enzyme	Buffer
BglII	NEB3
BamHI	NEB3 + BSA
EcoRI	Multicore buffer
HindIII	NEB2

Gel Electrophoresis

All gels were made with 0.7 g of agarose for every 100 ml of Tris-Acetate-EDTA buffer (TAE). The samples were loaded into the wells through the addition of loading dye, made with 30% glycerol and .25% bromophenol blue, with 1 µl of dye to 5 µl of sample, and the samples were electrophoresed at 60V. After electrophoresis, the gels were placed in a solution of ethidium bromide for staining, and viewed using UV light with a camera connected to a computer. The markers used for all gels are ΦX174 digested with HaeIII and λ DNA digested with HindIII, both obtained from Elizabeth Wyckoff.

Mating

The mating was used for the transfer of the pRES1 plasmid from SM10 (λ pir) into two *V. cholerae* strains, O395 and ALV101.

Overnight cultures of the SM10 clones containing pRES1 and the *V. cholerae* strains O395 and ALV101 were grown in 3 ml of L-broth, and kanamycin was used in the overnight cultures of the two pRES1 clones, while streptoymycin was used in the overnight cultures of the *V. cholerae* strains. 250 µl of each overnight culture was centrifuged in an Eppendorf tube and the supernatant was taken off, then the bacteria were resuspended in 5 µl of L broth. ALV101 and the SM10 strains containing pRES1 were placed together on one LB plate without antibiotics,

and the combinations of ALV101 and both SM10 strains containing pRES1, and O395 and both SM10 strains containing pRES1 were also spotted together on different LB plates. After 6 hours at 37°, the growth from these plates was put on plates containing both streptomycin and ampicillin (amp/strep) using a three phase streak for isolation. Eight single colonies were picked from each of these plates and streaked on another amp/strep plate, and these colonies were then grown up overnight in L-broth containing streptomycin, of which 3 ml was centrifuged at 13200 rpm and resuspended in 1 ml of TSB with 20% glycerol to store at -80° C.

Results:

Integration of the *feo* promoter into the PGOA1193 vector

Since the goal of this project was to ascertain when the Feo iron transport system is expressed, its promoter sequence was first integrated into the pGOA1193 vector, the vector that contains the resolvase. The *feo* 5' untranslated region containing the promoter sequence was amplified by taking Lou 15 DNA, or DNA acquired from the Lou 15 strain of *V. cholerae*, and by a PCR using Pfx on this DNA. The PCR was done using the Pfx PCR protocol outlined in the Materials and Methods section. The primers used for this PCR were Feo.908.Bgl and Feo.1852.Bgl. The agarose gel of the PCR product is shown below (Figure 2) with the PCR product in the first lane and the marker ΦX174 HaeIII in the second lane.

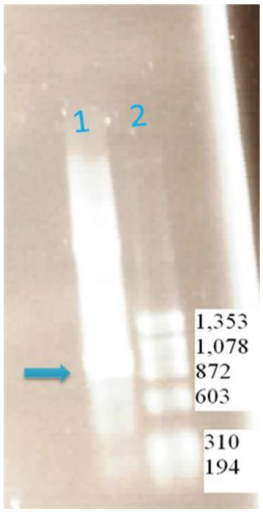

Figure 2: PCR of the *V. cholerae feo* untranslated region
Lane 1: *feo* PCR product
Lane 2: ΦX174 HaeIII

← product between 872 and 1,078 bp

Although the PCR product seems to be a large smear, there is a brighter area of the smear between 872 and 1,078 bp, which matches with the predicted size of the product from this PCR of 960 bp. Because the PCR fragment was not pure, I electrophoresed the entire PCR product overnight and did a gel extraction using the procedure outlined above. The gels shown below are that of the overnight gel (Figure 3) and the final gel extraction of the PCR product (Figure 4). Both are shown using the ΦX174 HaeIII ladder.

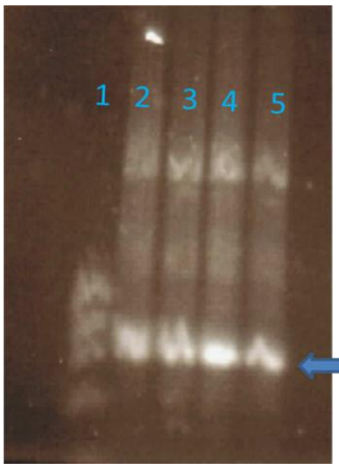

Figure 3: Overnight electrophoresis of *feo* PCR product
Lane 1: ΦX174 HaeIII
Lanes 2-5: *feo* PCR product

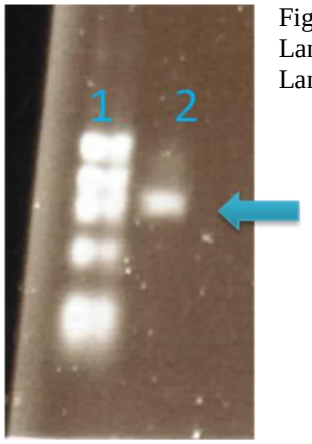

Figure 4: Gel electrophoresis of the purified *feo* PCR product
Lane 1: ΦX174 HaeIII
Lane2: gel purified *feo* PCR product

I then ligated the *feo* promoter, using the ligation procedure outlined earlier, into the pWKS30 vector that had been cut with SmaI. I then used this vector ligated with the *feo* promoter to transform competent SM10 cells made by Elizabeth Wyckoff using the heat-shock transformation protocol outlined in the Materials and Methods section. After growing colonies on LB plates with ampicillin, I grew 3.5 ml cultures of four of my clones in L-broth with ampicillin, and did another plasmid preparation with the same protocol as outlined before. When I examined the plasmid preparations of my clones against the pWKS30 vector that was cut with SmaI, I saw that the ligated vector was bigger (Figure 5).

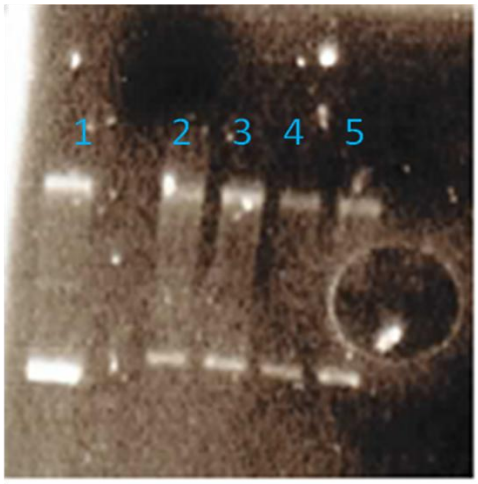

Figure 5: Gel of the plasmid preparations for pWKS30 vector ligated with the *feo* untranslated region
Lane 1: pWKS30 vector without ligated promoter
Lane 2: Plasmid preparation of pWKS30+ *feo* promoter clone 23
Lane 3: Plasmid preparation of pWKS30+ *feo* promoter clone 25
Lane 4: Plasmid preparation of pWKS30+ *feo* promoter clone 26
Lane 5: Plasmid preparation of pWKS30+ *feo* promoter clone 27

I then was able to take the *feo* promoter out of the pWKS30 vector using a BglII restriction enzyme digest, and ligate it into the pGOA1193 vector. This vector was obtained by doing a plasmid preparation on the strain containing pGOA1193 sent by Andrew Camilli at Tufts University. Both pGOA1193 and the pWKS30 vector have BglII sites, so this restriction enzyme was used to digest both the pGOA1193 vector so that it would be linearized, and to obtain the *feo* promoter fragment from the pWKS30 vector. The gel below (Figure 6) shows both the linearized pGOA1193 vector cut with BglII after it has been gel purified, using the Sigma gel purification procedure outline above (orange arrow), and the plasmid preparations of the pWKS30 vector ligated with the *feo* promoter (blue arrow).

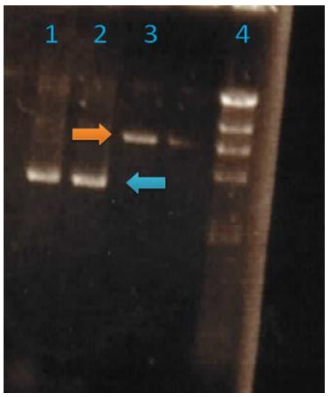

Figure 6: Gel of the plasmid preparations of the pWKS30 vector and the *feo* promoter, and the linearized pGOA1193 vector cut with BglII
Lane 1: Plasmid preparation of pWKS30+*feo* promoter clone 23
Lane 2: Plasmid preparation of pWKS30+*feo* promoter clone 25
Lane 3: pGOA1193 vector cut with BglII, gel purified
Lane 4: λ DNA- HindIII – marker

The next gel (Figure 7) shows the plasmid preparations of the two most successful ligations of the pWKS30 vector with the *feo* promoter, clones 23 and 25, after being digested with the BglII restriction enzyme.

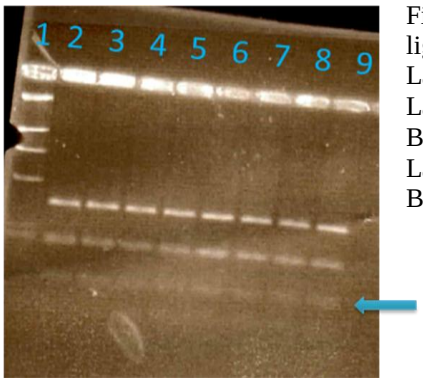

Figure 7: Overnight gel electrophoresis of the BglII digests of two vectors ligated with the *feo* promoter
Lane 1: λ DNA- HindIII – marker
Lanes 2-5: Plasmid preparation of pWKS30+*feo* promoter clone 23 cut with BglII
Lanes 6-9: Plasmid preparation of pWKS30+*feo* promoter clone 25 cut with BglII

The band shown by the blue arrow, migrating at less than 2,027 bp, was the predicted size for the *feo* promoter of 906 bp. I then used the same Sigma gel extraction kit to purify this band, and after purification, I electrophoresed the DNA out on another gel to check the size (Figure 8).

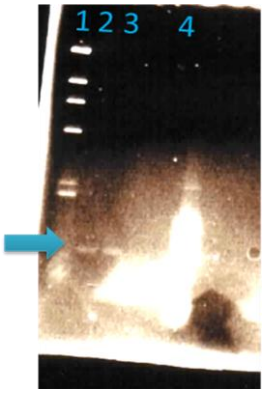

Figure 8: Gel of the purified *feo* promoter after having been cut out of the pWKS30 vector with BglII
Lane 1: λ DNA- HindIII – marker
Lane 2: Gel purification of the third band from the Plasmid preparation of pWKS30+*feo* promoter clone 23 cut with BglII
Lane 2: Gel purification of the third band from the Plasmid preparation of pWKS30+*feo* promoter clone 25 cut with BglII
Lane 4: ΦX174 HaeIII

I ligated the *feo* promoter, the correctly sized band cut from the plasmid preparation, into the cut pGOA1193 vector using the same ligation protocol as outlined in the Materials and Methods section. I then used this vector ligated with the *feo* promoter to do a heat-shock transformation of SM10 *E. coli* cells as outlined earlier. I then picked colonies that grew on an ampicillin plate and transferred them onto another ampicillin plate, and I did a Pfx PCR using the same procedure with the Feo.908.Bgl and Feo.1852.rev.Bgl vectors to see whether the clones had the *feo*

promoter ligated into the PGOA1193 vector. The results of this PCR are shown below (Figure 9).

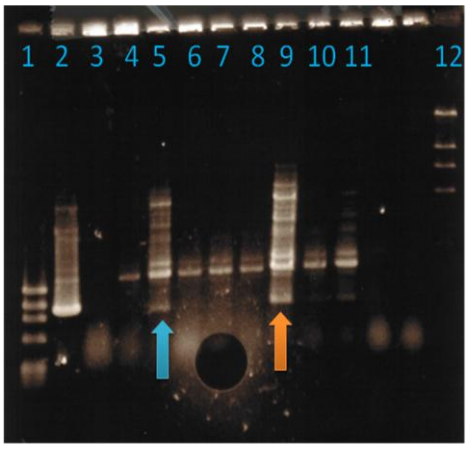

Figure 9: PCR results of using the *feo* BglII primers on the ligation of the *feo* promoter into the pGOA1193 vector
Lane 1: λ DNA- HindIII - marker
Lane 2: pWKS30 with *feo* promoter (positive control)
Lane 3: Empty pGOA1193 vector (negative control)
Lane 4: pGOA1193with *feo* promoter clone 1
Lane 5: pGOA1193with *feo* promoter clone 2
Lane 6: pGOA1193with *feo* promoter clone 3
Lane 7: pGOA1193with *feo* promoter clone 4
Lane 8: PGOA1193 with *feo* promoter clone 5
Lane 9: pGOA1193with *feo* promoter clone 6
Lane 10: pGOA1193with *feo* promoter clone 7
Lane 11: pGOA1193with *feo* promoter clone 8
Lane 12: ΦX174 HaeIII

The most promising clones were those which looked most like the positive control, clones 2 and 6. I then grew up overnight plates of these clones and stored the bacteria with TSB and 20% glycerol at -80° C. These are the clones of pGOA1193 ligated with the *feo* promoter that I used in all subsequent experiments.

Transformation of DH5α (λ pir) cells with pRES1

The first step in achieving a successful transformation was to test the pRES and pRES1 vectors, to be sure that the vectors showed the correct banding pattern when cut with different restriction enzymes, and to be sure that the vectors were the correct size. I performed a plasmid preparation using the Sigma minipreparation procedure outlined in the Materials and Methods section, to obtain the plasmid DNA. I first digested these vectors with the digestion enzymes BamHI, EcoRI and HindIII to check the banding patterns, to see if these patterns matched those which would be expected after analyzing the sequence of the plasmids sent by Andrew Camilli. The gel is shown in Figure 10.

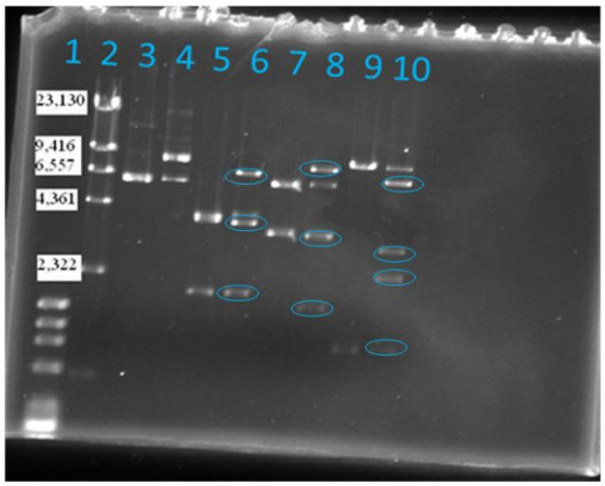

Figure 10: Gel of the pRES vector digested with different restriction enzymes
Lane 1: ΦX174 HaeIII-marker
Lane 2: λ DNA- HindIII – marker
Lane 3: Undigested pRES vector
Lane 4: Undigested pRES1 vector
Lane 5: pRES vector digested with BamHI
Lane 6: pRES1 vector digested with BamHI
Lane 7: pRES vector digested with EcoRI
Lane 8: pRES1 vector digested with EcoRI
Lane 9: pRES vector digested with HindIII
Lane 10: pRES1 vector digested with HindIII

Since pRES and pRES1 should have identical sequences except for a 1 base change difference at the resolution crossover site, they should not have different banding patterns. In the first two lanes, those that contain the undigested vector, the pRES vector is running at around 6,500 bp and the pRES1 vector has bands at both that length and at around 9, 416 bp. Unlinearized vectors usually run faster than linearized vectors, due to supercoiling, but as the vector is supposed to be 12,627 bp, the band running at 9,416 bp seems more likely to be accurate. From the gel, pRES1 has both bands that migrated the same distance as the pRES bands, and bands that are different. This follows from fact that the undigested pRES1 vector shows two different bands, one the same as the pRES vector, and one that is larger. After analyzing these bands, I found that the lanes containing pRES1 contained the correct size bands after cutting: 7109, 3684 and 1724 for BamHI, 7785, 3261 and 1471 for EcoRI, and 6640, 2854, 2109 and 914 for HindIII. These correctly sized fragments are shown by the blue circled banding pattern in Figure 10. The pRES lanes contained some of these bands, but not all. This led to a hypothesis that the pRES plasmid was somehow missing the sequence flanked by the resolvase that would lead to kanamycin resistance and sucrose sensitivity. After seeing this phenomenon, I then made sure of the size of

the vectors by digesting with NotI, a digest enzyme which only cuts in one place on the vector (Figure 11).

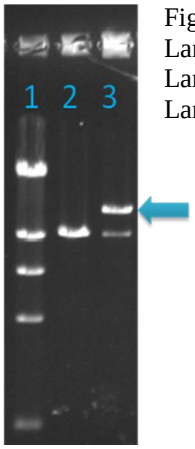

Figure 11: Gel of pRES and pRES1 digested with NotI
Lane 1: λ DNA- HindIII – marker
Lane 2: pRES digested with NotI
Lane 3: pRES1 digested with NotI

Correct sized fragment for an opened pRES1 vector-12,671 bp

The pRES vector digested with NotI runs at around 9,412 bp, while the pRES1 vector digested with NotI runs at both 9,412 bp as well as around 12,500 bp. This is consistent with the hypothesis that the *neo* and *sacB* genes, leading to resistance to kanamycin and sensitivity to sucrose, were missing from the plasmid, since these genes are about 3 kb together. The loss of these genes from the plasmid is possible if the plasmid were to spontaneously resolve these sequences, which could occur depending on the genetic background of the bacteria in which they were maintained. To make sure that this is what happened, I grew the strains containing pRES and pRES1 on plates containing ampicillin and plates containing kanamycin, and only pRES1 grew on both plates, demonstrating that some of the bacteria still hold the complete pRES1 plasmid. I then transformed DH5α (λ pir) *E. coli* with my plasmid preps of pRES and pRES1, using the heat shock method, and grew the resulting transformed cells on media containing Kan. The DH5α (λ pir) strain does not contain kanamycin resistance, so the only way these bacteria would be kanamycin resistant is by taking up the plasmid. Only the bacteria transformed with pRES1 gained kanamycin resistance, so I picked eight single colonies to grow up on a kanamycin plate, and then made overnight cultures of these colonies in 3 ml of L-broth with

kanamycin to perform a plasmid preparation. I then performed a Sigma plasmid preparation using the same procedure to obtain the plasmid DNA from these bacteria. The gel is shown in Figure 12.

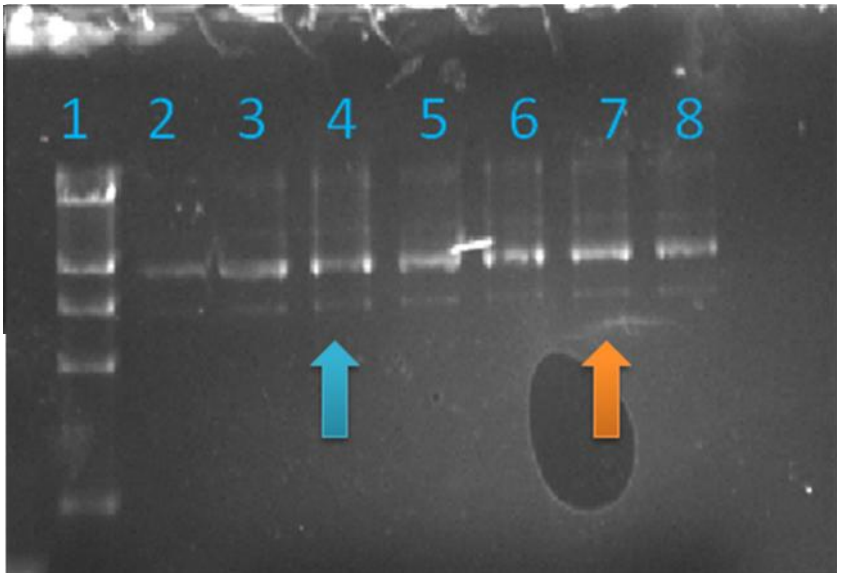

Figure 12: Gel of the plasmid preparation of the transformation of DH5α (λ pir) with pRES1
Lane 1: λ DNA- HindIII – marker
Lane 2: DH5α (λ pir) and pRES1 vector clone 1
Lane 3: DH5α (λ pir) and pRES1 vector clone 2
Lane 4: DH5α (λ pir) and pRES1 vector clone 3
Lane 5: DH5α (λ pir) and pRES1 vector clone 4
Lane 6: DH5α (λ pir) and pRES1 vector clone 6
Lane 7: DH5α (λ pir) and pRES1 vector clone 7
Lane 8: DH5α (λ pir) and pRES1 vector clone 8

From this gel I determined that clone 3 (blue arrow) and clone 7 (orange arrow) were the best candidates for further use, since they were brighter than most of the other bands. All bands migrated at around 9,416 bp, which is the distance the uncut pRES1 vector which contained the correct *neo* and *sacB* sequences migrated. These clones will be used in all subsequent experiments.

Transformation of SM10 cells with the pRES1 plasmids

After obtaining the plasmids from my transformations of DH5α (λ pir) with pRES1, I used these plasmids to transform SM10 *E. coli* cells. The strain DH5α (λ pir) does not contain the correct genes for conjugation with another bacteria, so I must transfer these plasmids to SM10, which can successfully mate with other strains of bacteria. This procedure was done using the same transformation procedure as outlined in the Materials and Methods section. The strains

were grown on plates using ampicillin, because the pRES1 vector confers this resistance to the bacteria, and the SM10 strain is already kanamycin resistant. This research is currently ongoing.

Mating of SM10 cells with pRES1 into O395 and Alv101 *V. cholerae*

After obtaining the SM10 cells containing the pRES1 plasmid, I will mate these bacteria into the *V. cholerae* strains O395 and Alv101 using the protocol for mating outlined in the Materials and Methods section. From the overnight culture of the 8 different clones, 100 µl will be centrifuged to form a pellet and resuspended in 1 ml of water, and this will be used in PCR reactions using the LacZ and Rivet primers, using the Taq PCR procedure outlined in the materials and methods section..

The most promising clones, which contain the correct pRES sequence, will be grown up overnight in streptomycin, and then 40 µl of this growth will be transferred into L-broth containing kanamycin, to grow up the strains that correctly incorporated the plasmid. These strains will be used in subsequent experiments.

Discussion:

The creation of a strain that could be used to test for the transcription of the *feo* gene will be a very useful tool in ascertaining when this iron transport system is expressed. There were a few problems in making this strain, however. The incorporation of the *feo* promoter into the pGOA1193 vector went very smoothly, but the transformation using the pRES and pRES1 strains did not happen as easily. I first attempted to transform SM10 λ-pir competent cells with these plasmids, but as these cells are kanamycin resistant already, I did not select for kanamycin resistance donated from the plasmid. This meant that the SM10 cells were transformed with the plasmids lacking the *neo* and *sacB* sequences, as shown in the results section. This explains why when I attempted to undertake the mating with these competent cells and the *V. cholerae* strains none of the strains became kanamycin resistant, since the plasmid within the SM10 cells did not contain the *neo* gene. This was resolved, however, after undertaking the experiments outlined in the results section.

There are many experiments that should be completed using the successful strain created using the RIVET vector integrated with the *feo* promoter. First the strain must be tested to see how often it resolves when outside a host, by growing up the strain in L broth and then testing how well it retains kanamycin resistance and sucrose sensitivity. After these initial tests, the strain can be observed in a variety of media containing different amounts of iron and different types of iron, such as ferrous and ferric, to further ascertain when this iron transport mechanism is active. The strain can then be used to infect infant mice, and the strain can then be tested to find whether the *feo* promoter is active during infection with the *V. cholerae* bacteria. The use of this strain to find when the *feo* gene is expressed could potentially lead to great advances in the understanding of the different iron mechanisms which *V. cholerae* uses. The pGOA1193 vector

can also be integrated with other iron transport gene promoters as well as other promoter sequences in the *V. cholerae* gene library, so that further analysis of when this bacteria expresses its various genes can be performed.

Citations:

C. Andrus, Walter M., Crosa J., Payne S., Synthesis of siderophores by pathogenic Vibrio species. *Curr Microbiol,* 9. pp 209–214. (1983).

J. Butterton, Boyko S., and Calderwood S., Use of the *Vibrio cholerae irgA* gene as a locus for insertion and expression of heterologous antigens in cholera vaccine strains. *Vaccine,* 11**.** pp1327-1335. (1993)

M. Cartron, Maddocks S., Gillingham P., Craven C., and Andrews S., Feo - transport of ferrous iron into bacteria. *Biometals*, 19(2). pp 143-157. (2006).

S. Faruque, Albert M., and Mekalanos J., Epidemiology, Genetics, and Ecology of Toxigenic *Vibrio cholerae*. *Microbiology and Molecular Biology Reviews,* 62(4). pp 1301-1314. (1998).

D. Henderson, and Payne S., *Vibrio cholerae* iron transport systems: roles of heme and siderophore iron transport in virulence and identification of a gene associated with multiple iron transport systems**.** *Infection and Immunity*, 62(11). pp 5120-5125. (1994)

A. Mey and Payne S., Haem utilization in *Vibrio cholerae* involves multiple TonB-dependent haem receptors. *Mol. Microbiol*, 42. pp 835-849. (2001)

P. Murray, Rosenthal S., Pfaller M., *Medical Microbiology*. Elsevier Inc. (2005)

J. Neilands., Siderophores: structure and function of microbial iron transport compounds. *J. Biol. Chem*, 270. pp 26723-26726. (1995)

D. Occhino, Wyckoff E., Henderson D., Wrona T., Payne S., *Vibrio cholerae* iron transport: haem transport genes are linked to one of two sets of *tonB, exbB, exbD* genes. *Mol Microbiol,* 29. pp 1493–1507. (1998).

C. Osorio, Crawford J., Michalski J., Martinez-Wilson H., Kaper J., and Camilli A. Second-generation recombination-based in vivo expression technology for large scale screening for *Vibrio cholerae* genes induced during infection of the mouse small intestine. *Infection and Immunity*, 73. pp 972-980. (2005)

K. Postle. TonB and the Gram-negative dilemma. *Mol. Microbiol*, 4(12). pp. 2019-2025 (1990)

S. Payne. Iron and virulence in *Shigella*. *Mol. Microbiol*, 3(9). pp 1301–1306. (1989)

S. Payne, and Finkelstein R., The critical role of iron in host-bacterial interactions. *J Clin Invest*, 61(6). pp 1428–1440. (1978)

S. Payne, Wyckoff, E., Murphy, E., Oglesby, A., Boulette M., and Davies N. Iron and pathogenesis of *Shigella*: iron acquisition in the intracellular environment. *Biometals*. 19. pp 173-180. (2006)

J. Reidl, and Klose K., *Vibrio cholerae* and cholera: out of the water and into the host. *FEMS Microbiol.* 26. pp 125-139. (2002)

D. Sack, Sack R., Chaignat C., Getting serious about cholera. *N Engl J Med*, 355. pp 649–651. (2006).

S. Seliger, Mey A., Valle A., and Payne S. The two TonB systems of *Vibrio cholerae*: redundant and specific functions. *Mol. Microbiol*, 39(3). pp. 801-812 (2001)

P. Watnick, Butterton J., Calderwood S., The interaction of the *Vibrio cholerae* transcription factors, Fur and IrgB, with the overlapping promoters of two virulence genes, *irgA* and *irgB*. *Gene*, 209. pp 65–70 (1998)

E. Wyckoff, Mey A., and Payne S., Iron acquisition in *Vibrio cholerae*. *Biometals*, 20. pp 405-416. (2007)

E. Wyckoff, Mey A., Leimbach A., Fisher C., and Payne S., Characterization of ferric and ferrous iron transport systems in *Vibrio cholerae*. J Bacteriol, 188(18). pp. 6515–6523. (2006)

Appendix 1: pRES sequence

pRES (pAC1311): 12,627 bp. VcvchoR' 105-1, VcchrR 1389-247, VclacZ' 1712-3161, res 3189-3302, neo 3424-4239, sacB 4748-6169, res 6303-6416, Vc'lacZ 6649-8190, bla 9112-9972, traI 10982-10213. **Note that the genes beginning with "Vc" are *V. cholerae* sequences for recombining the cassette into the Vc lacZ gene.**

GATATCAGGCAGCTCTACCCGAATTTTCTCGAGCATATTGGTTAAATAGGTATTTAGCGACCCC
ACCGATGGGTAATAGCGAAACTGACCGACTTTGGGTTGCATCTGCATATCACCGCCCTTCCTTG
CTAGCTACTCATTAAGTTTAGACAAGGTTTTGTGGATTGCCGCATTAGGGCGGTGCTTTTCTAT
GTCTGTTTTACTACTCGAGGTCTGTTTTACTACTCGATCACATCCCGTTGCTTTTTAGATTAAG
CCGCTGAGCACACCAAATAACAAGGTGCTCCCCACCATGCCAATTAAAGGAAGTCGCAAGACTT
TAAACAAGAAGAAGCCGACTAACAGGGCTGCCCAATCGAGTGGGGCATGTACTGTGCTACTCCA
AACGGGTTGATATAAGGCGGCGAGCAGTAAACCGACCACACAGGCGTTTACTCCTTGCATGGCT
CCTGCAACTTTCGGGCGCTGAGCCAATGCCGACCAGTTTTTGAGTACGCCGAGTAGCAATAAAA
AACCGGGGAGAAACACGGCTAAGGTTGCGATTAAAGCGCCAACCACGGGCATCTCTGGTGTCAG
CACATAGCCAAGGTAAGTGGCTAAGGTAAACATTGGGCCCGGCACCGCTTGTGCTGCCGCGTAG
CCTGTCAGAAATTGATCGGTCGATAGGCTGTCACCTAAGGCATTTTGCAGCAAGGGAAGAACCA
CATGACCGCCACCAAACACTAAGCTTCCTGCTTGGAAGAAATGGCCAAACAACTCAACCAGCGG
TGAGCTTGCAGCCAAGAGCGGCAGCCCGAGTAATAAACTTGCAAAAAGTAGCAGTGAAAGCCAA
GAGGGGCTGAACGTGGTTGTCGAAAATGACTGTTGTGGTGCCAAGCGGGCTTGGCCAACGAAAG
CGGCGATCAGAAGGACCGCAAACTGAGTGATGAGACCGGGAGCTAACGTAATCGCAACCGCAGT
CAGAACACATAACCCTGCAGTAAGGCGTTGCTGACAAAAATTGCGATACATGGTTAAACAAGCA
TCGGCCACCACTATGATCGCGAGTAGCTTCAAGCCATGAATCATTTGTTCAAACAGTGGGGTAT
CAAGGAGATGGCTGCTTAGCCCCGCGAGCAGCAGCATGATGAGCACGGAGGGAAGGGTAAAACC
GAGAAACGCTGCCCAAGCGCCAGCCAGACCACCACGATGATAACCAATCGCAAAACCAACTTGG
CTAGAACCGGGGCCAGGAAGGAACTGGCTCAGTGCCACAAATTGTGCATATTCTTGCTCGCTAA
CCCAGCGTAATTTCTCAACAAAGGTGTGGCGAAAGTAACCTATATGTGCGGCTGGCCCACCAAA
ACTCACCCATCCGAGAGCAAAAAAAGTTCTAAAAATCGTTAGCATAATGATCTGAAGTCATCCG
TAATCAATGGAAGGTCAACATCCGTAGGAGCATAGGTTATGGAGAGTCAAAGCGCAGAACAACT
CCGAATGTGTAAAAAATTACACCATTATGATGGCAATCGTATGAATCGATTCAGAAATAGAAAA
ATTGGGTCAATATCGACCTCTATTTAAATTGTGGAAACGTTTACACAATTGGTGAGTGGTTCAC
AGAATCGGTGTTTGAAAGTTTGTTAGACTTTTTTGCATCTGCAGCATGTCATCATTCCTATTCA
AAGCTGCGAATCTTATTGAATGACTTCTTTACTCCTCGGCTTGAGGGATGTACGCCGTAGAGCA
AAGGCGTTATTGGCTTGTTGCTTCGCGTATTCAAGAGGATCTTATGCGCAACTTCTCCGATATT
CTTCTTAGCCAAGATTGGCAAAACCCGCACATCGTTAAATGGCACTGCCGTACACCCCATGTTC
CTTTGCACAGTTATCGCACTGAGCAGGAGGCTCGTTTGGATGTTGGGGGGAATCGCCAATCTCT
AAATGGTCAGTGGCGGTTTGCTCTGTTTGAGAAGCCAGAAGCGGTTGAGCCTGCGGTGATAGAC
CCGGATTTCGATGATAGCGCTTGGGCGCACATTCCTGTACCGAGTAACTGGCAGATGCAAGGCT
TTGATAAGCCGATTTACACCAATATCCAATATCCATTTGCGGATCGGCCGCCTTACGTGCCGCA
AGATAATCCAACCGGCTGTTATCGCCACCGTTTTACACTGGAAAAACAAGCGCTAACCGAGTCC
ATTCGCATTGTATTTGATGGGGTCAATTCGGCATTTCATCTGTGGTGCAATGGTCATTGGGTCG
GTTATTCGCAAGATAGCCGCTTGCCTGCCGAGTTTGAGTTAACCCCTTATCTACAAGAGGGTGA
AAACCTGTTGGTGGCCATGGTGCTGCGCTGGTCTGATGGCTCTTATTTGGAAGACCAAGATATG
TGGTGGCTGAGTGGCATCTTTCGCGATGTGTATCTCTACCGCAAGCCGATACTCGCGATTGAAG
ATTTTTTTATCCGCACTGAATTAGATGCGCTTTATCAACACGCTGAATTGCGAGTAGAAACACG
CTTAAGCCAAGTGACTCGCCATCATCAAGTGCAAGTGGCTTTATTCGATGCACAAGGTGAATGC

GTGGCGCGTTCACAAGCCTTACATACAGGCCAGCGTGTAGTGGATGAAAAAGGAGCATGGCACG
ATAAAACCGAACACAGTTTAGCGATTTGCTCTCCGACACTGTGGAGTGATGAAGCGCCTTATCT
TTACCGCTGCGTGATCTGTTTGCTTGATGAAGATGGCGCGCCGATTGAGTTTGAAAGTGCAGCA
GTGGGTTTTCGCAAAGTAGAAATCACTCAGGGACTACTGAAGCTCAATGGTCAGCCCTTGTTGA
TCCGCGGGGTGAACCGTCATGAACATCATCCCGAATTGGGGCATGTGATGGATGAAGCAAGCAT
GCGCCGCGATATTGAATTGATGAAACAGCATAATTTCAATGCGGTGCGTACCGCCCATTACCCC
AATCATCCGCGTTGGTACGAACTGTGTGATGAGTACGGTTTGTATGTGGTGGATGAGGCCAATC
TCGAAACCCACGGCCAATTTCCGATGAGCCGACTTTCCAATGATCCACAATGGGTGAATGCCTA
TTTGCAGCGCATGATTGGCATGGTGGAGCGCGATAAAAACCACCCTTGTGTGATCATTTGGTCG
CTCGGCAATGAATCGGGGATTGAATTAACCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCG
CGGTGGCGGCCGCTCTAGCCCGGGCGGATCCCCCGGGCTGCAGGAATTCCGTCCGAAATATTAT
AAATTATCGCACACATAAAAACAGTGCTGTTAATGTGTCTATTAAATCGATTTTTTGTTATAAC
AGACACTGCTTGTCCGATATTTGATTTAGGATACACTCGAGTCTAGAAAAGGTTAGGAATACGG
TTAGCCATTTGCCTGCTTTTATATAGTTCATATGGGATTCACCTTTATGTTGATAAGAAATAAA
AGAAAATGCCAATAGGATATCGGCATTTTCTTTTGCGTTTTTATTTGTTAACTGTTAATTGTCC
TTGTTCAAGGATGCTGTCTTTGACAACAGATGTTTTCTTGCCTTTGATGTTCAGCAGGAAGCTT
GGCGCAAACGTTGATTGTTTGTCTGCGTAGAATCCTCTGTTTGTCATATAGCTTGTAATCACGA
CATTGTTTCCTTTCGCTTGAGGTACAGCGAAGTGTGAGTAAGTAAAGGTTACATCGTTAGGATC
AAGATCCATTTTTAACACAAGGCCAGTTTTGTTCAGCGGCTTGTATGGGCCAGTTAAAGAATTA
GAAACATAACCAAGCATGTAAATATCGTTAGACGTAATGCCGTCAATCGTCATTTTTGATCCGC
GGGAGTCAGTGAACAGGTACCATTTGCCGTTCATTTTAAAGACGTTCGCGCGTTCAATTTCATC
TGTTACTGTGTTAGATGCAATCAGCGGTTTCATCACTTTTTTCAGTGTGTAATCATCGTTTAGC
TCAATCATACCGAGAGCGCCGTTTGCTAACTCAGCCGTGCGTTTTTTATCGCTTTGCAGAAGTT
TTTGACTTTCTTGACGGAAGAATGATGTGCTTTTGCCATAGTATGCTTTGTTAAATAAAGATTC
TTCGCCTTGGTAGCCATCTTCAGTTCCAGTGTTTGCTTCAAATACTAAGTATTTGTGGCCTTTA
TCTTCTACGTAGTGAGGATCTCTCAGCGTATGGTTGTCGCCTGAGCTGTAGTTGCCTTCATCGA
TGAACTGCTGTACATTTTGATACGTTTTTCCGTCACCGTCAAAGATTGATTTATAATCCTCTAC
ACCGTTGATGTTCAAAGAGCTGTCTGATGCTGATACGTTAACTTGTGCAGTTGTCAGTGTTTGT
TTGCCGTAATGTTTACCGGAGAAATCAGTGTAGAATAAACGGATTTTTCCGTCAGATGTAAATG
TGGCTGAACCTGACCATTCTTGTGTTTGGTCTTTTAGGATAGAATCATTTGCATCGAATTTGTC
GCTGTCTTTAAAGACGCGGCCAGCGTTTTTCCAGCTGTCAATAGAAGTTTCGCCGACTTTTTGA
TAGAACATGTAAATCGATGTGTCATCCGCATTTTTAGGATCTCCGGCTAATGCAAAGACGATGT
GGTAGCCGTGATAGTTTGCGACAGTGCCGTCAGCGTTTTGTAATGGCCAGCTGTCCCAAACGTC
CAGGCCTTTTGCAGAAGAGATATTTTTAATTGTGGACGAATCGAATTCAGGAACTTGATATTTT
TCATTTTTTTGCTGTTCAGGGATTTGCAGCATATCATGGCGTGTAATATGGGAAATGCCGTATG
TTTCCTTATATGGCTTTTGGTTCGTTTCTTTCGCAAACGCTTGAGTTGCGCCTCCTGCCAGCAG
TGCGGTAGTAAAGGTTAATACTGTTGCTTGTTTTGCAAACTTTTTGATGTTCATCGTTCATGTC
TCCTTTTTTATGTACTGTGTTAGCGGTCTGCTTCTTCCAGCCCTCCTGTTTGAAGATGGCAAGT
TAGTTACGCACAATAAAAAAAGACCTAAAATATGTAAGGGGTGACGCCAAAGTATACACTTTGC
CCTTTACACATTTTAGGTCTTGCCTGCTTTATCAGTAACAAACCCGCGCGATTTACTTTTCGAC
CTCATTCTATTAGACTCTCGTTTGGATTGCAACTGGTCTATTTTCCTCTTTTGTTTGATCGAAA
ATCATAAAAGGATTTGCAGACTACGGGCCTAAAGAACTAAAAAATCTATCTGTTTCTTTTCATT
CTCTGTATTTTTTATAGTTTCTGTTGCATGGGCATAAAGTTGCCTTTTTAATCACAATTCAGAA
AATATCATAATATCTCATTTCACTAAATAATAGTGAACGGCAGGTATATGTGATGGGTTAAAAA
GGATCAGATCTAATGCTCTGCCAGTGTTACAACCAATTAACCAATTCTGATTAGAAAAACTCAT
CGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAG
CCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTAT

CGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAA
GGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGCTTATG
CATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCA
ACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAG
GACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATT
TTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCGCAGTGGTG
AGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCG
TCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTT
CAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCG
ACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCC
TCGAGCAAGACGTTTCCCGTTGAATATGGCTCATTCTAGACGTCCGAAATATTATAAATTATCG
CACACATAAAAACAGTGCTGTTAATGTGTCTATTAAATCGATTTTTTGTTATAACAGACACTGC
TTGTCCGATATTTGATTTAGGATACAGGTACCCAATTCGCCCTATAGTGAGTCGTATTTACCAA
TCATCACGCCATGTATCAGTGGACGAAACAGCGCGACCCATCGCGTCCTGTGCAATACGAAGGG
GGCGGCGCTAATACGGCGGCGACCGATATTGTTTGCCCGATGTATGCGCGGGTCGATCAGCATC
AGCCACATCTGCGGTTCCAAAATATGCGCTGAAAAATTGGATCAGTTTGCCGCAGGAAAACCGC
CCCCTCATCTTGTGTGAATATGCTCATGCGATGGGCAACAGCTTGGGCGCGTTTTATAAATACT
GGCAGGCGTTTCGTGAGTTTCCTCGTCTGCAAGGTGGCTTTATTTGGGATTGGGTCGATCAGGG
CATTTCCAAATGGGATAGCGAGGGGCGCCACTATTGGGGCTATGGCGGTGATTTTGGCGATACG
ATTAACGATCGCCAATTCTGCATAAACGGTTTGCTGTTCCCAGATCGCACGCCGCATCCGGCAT
TACATGAAGTCAAAAAAGTCCAGCAGCCGTACCAGTTTTCGTTGAGCTATCCCAAGCTCACCAT
TCACAATGAGCGCTTGTTTGCAGCGCTGCCGCTGGAGCTGGTAGTTAGTGTGCTATGCGATGGG
CAAGAGATTAAGCAAGAACGTCTGCCGCTTGATATTGCGCCGCGCGGCACAATCACGCTGGATT
TAGCGTCGCTGCCAATGTTGCCAGAGCATGAATACCACCTCAATGCAGTCTTATTGTGTCGTGA
GGATCAGCCATGGTCTAACGCGGGGCACTGCATCGCTAGTGAGCAGTGGTGTTTGCAGCCACGA
AGAAGCATGTTACCTAAAATCACACACGCTCCGCTGCCTCAATGGCAGCAAGATGGAGATAAGG
TGCGCATCGAGGCGGCCAATCAGCAATGGCAGTTTAACCGCCAAACTGGGCTATTGGAGCAGTG
GTGGCAAAATGGTCAGCCCGTATTGAGTGAACCGCTGCGCGATAACTTTTACCGCGCGGTGCTG
GATAACGATATTGGTACTAGCGAAGCGCAGCATCTTGACCCGAACAGCTGGATCGCACGTTGGC
ATGCGGCGGGCTTAGATAAGCTGCGTGTGGAATGTGACGATCTTCGCGTCACCACCTTGAACGA
GAGTGTCGAAGTGGTGATCGATGTCGCCCATTACCATCAGCAAGCGTTAGCGCTTCGTACCCGT
TGGCGTTACCAAATCTTCGGTGATGCGCGGGTAGAACTGAATGTTGAGGTGATGCTGTGTTCTG
ATTTACCGCCGCTGCCAAGAGTGGGGTTAACGCTCGCATTACCAGTGGCAGAAAACCCAGTGTC
TTGGTTTGGTCGCGGGCCGCATGAGAATTATCCGGATCGTTTGCAATCGGCGCATGTGGGGCGA
TACACCGCCACGGTGGATGAGCTGCATACACCGTACATTTTCCCGAGCGAAAATGGTTTGCGTT
GTGATACTCGCCAGCTACAAGTGGGCGCTTTGGTGGTGGAAGGGCATTTTCACTTCTCGCTCAG
TCGCTACTCACAAACGATGTTGGATAAAGCCAAACACAGCAACGAGTTGGTGGCGGGCGATAAG
TGGTATCTCAATCTGGATGCGCAGCATATGGGCGTGGGCGGCGATGATTCGTGGAGCCAAAGTG
TGCACCCTGAATTTTTGCTCACTCAGCCGCACTATCAGTATCAGCTCACCTTACGTGTGAAAGC
GTCATCCCCACAATAAGCCAGAGAGCCTTAAGGCTCTCTTTTTTGTGCCACGCTGTCGCGGCGA
ACGATGGTGGGCGAAAAGATCATCGGATCATTTTCACGCGGTGCATCGGTTGCCAAAGCAAGAG
CCAACTGAGTGGCTTTTTCTGCCATCAGTTGGATTGGGTAACGCACGGTCGTCAGGCGTGGGCG
TAAGTAGCGCGAAATCAGGGCATCGTCAAAACCAATCACTGAGACATGTTCAGGAACCGAATGC
CCATTTTCCTCAAGTACCAACAGGGCTCCAGCGGCCATATTGTCGTTGTAAGCGACCACTGCGG
TAAAAGGCAGTGATTTCACCAGTAGGTTAGTCATCGCCTGTTCGCCGCCTTCACTGTCTGGGCT
GGCTTTTTCAATATAGCTGCTGCTTAGGGTGATGCCATGTTCGTTCAAGGCTTGTTGATAGCCC
GCAATACGCTGATCGGCATCCTCAATTTGATGAGAGGAGCTGATGCAAGCAATGTTTCTATGTC

CTTGACGGATGAGAAAATCGGTTGCTAAATAAGCACCTTTCTGGTTGTCGAGTGAAATACAGCG
ATCGGCGAGCTGAGGAATATGGCGATTGATCAACACCAGCGTTTTTACCTCGTTGGCGTACTCA
ATCAGCTCTTCATTGGGAAGCGCTTTGGAGTGGATGACCAGTGCATCACAGCGGCTGTTGATGA
GCAGTTCCAGCGCGCGGCGCTCTTCTTCCGCTCGGTGGTAGCCGTTGCCGATCAGCAAATGTTT
TCCTTCGCGGTGGGCGACAGTGTCGACCTGCAGATCTGCAGGTGGCACTTTTCGGGGAAATGTG
CGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAAT
AACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGT
CGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTG
AAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACA
GCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGT
TCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATA
CACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCA
TGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACT
TCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTA
ACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCA
CGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGC
TTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCG
GCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTA
TCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAG
TCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCAT
TGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAAT
TTTGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTCCGCTA
TCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGAC
GGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTG
TCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGCAAGGAGATGGCGCCCAACAGTCC
CCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGA
GCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGG
TGATGCCGGCCACGATGCGTCCGGCGTAGAGGATCCAGCCGACCAGGCTTTCCACGCCCGCGTG
CCGCTCCATGTCGTTCGCGCGGTTCTCGGAAACGCGCTGCCGCGTTTCGTGATTGTCACGCTCA
AGCCCGTAGTCCCGTTCGAGCGTCGCGCAGAGGTCAGCGAGGGCGCGGTAGGCCCGATACGGCT
CATGGATGGTGTTTCGGGTCGGGTGAATCTTGTTGATGGCGATATGGATGTGCAGGTTGTCGGT
GTCGTGATGCACGGCACTGACGCGCTGATGCTCGGCGAAGCCAAGCCCAGCGCAGATGCGGTCC
TCAATCGCGCGCAACGTCTCCGCGTCGGGCTTCTCTCCCGCGCGGAAGCTAACCAGCAGGTGAT
AGGTCTTGTCGGCCTCGGAACGGGTGTTGCCGTGCTGGGTCGCCATCACCTCGGCCATGACAGC
GGGCAGGGTGTTTGCCTCGCAGTTCGTGACGCGCACGTGACCCAGGCGCTCGGTCTTGCCTTGC
TCGTCGGTGATGTACTTCACCAGCTCCGCGAAGTCGCTCTTCTTGATGGAGCGCATGGGGACGT
GCTTGGCAATCACGCGCACCCCCCGGCCGTTTTAGCGGCTAAAAAAGTCATGGCTCTGCCCTCG
GGCGGACCACGCCCATCATGACCTTGCCAAGCTCGTCCTGCTTCTCTTCGATCTTCGCCAGCAG
GGCGAGGATCGTGGCATCACCGAACCGCGCCGTGCGCGGGTCGTCGGTGAGCCAGAGTTTCAGC
AGGCCGCCCAGGCGGCCCAGGTCGCCATTGATGCGGGCCAGCTCGCGGACGTGCTCATAGTCCA
CGACGCCCGTGATTTTGTAGCCCTGGCCGACGGCCAGCAGGTAGGCCGACAGGCTCATGCCGGC
CGCCGCCGCCTTTTCCTCAATCGCTCTTCGTTCGTCTGGAAGGCAGTACACCTTGATAGGTGGG
CTGCCCTTCCTGGTTGGCTTGGTTTCATCAGCCATCCGCTTGCCCTCATCTGTTACGCCGGCGG
TAGCCGGCCAGCCTCGCAGAGCAGGATTCCCGTTGAGCACCGCCAGGTGCGAATAAGGGACAGT
GAAGAAGGAACACCCGCTCGCGGGTGGGCCTACTTCACCTATCCTGCCCGGCTGACGCCGTTGG
ATACACCAAGGAAAGTCTACACGAACCCTTTGGCAAAATCCTGTATATCGTGCGAAAAAGGATG
GATATACCGAAAAAATCGCTATAATGACCCCGAAGCAGGGTTATGCAGCGGAAAAGCGCTGCTT

CCCTGCTGTTTTGTGGAATATCTACCGACTGGAAACAGGCAAATGCAGGAAATTACTGAACTGA
GGGGACAGGCGAGAGACGATGCCAAAGAGCTACACCGACGAGCTGGCCGAGTGGGTTGAATCCC
GCGCGGCCAAGAAGCGCCGGCGTGATGAGGCTGCGGTTGCGTTCCTGGCGGTGAGGGCGGATGT
CGAGGCGGCGTTAGCGTCCGGCTATGCGCTCGTCACCATTTGGGAGCACATGCGGGAAACGGGG
AAGGTCAAGTTCTCCTACGAGACGTTCCGCTCGCACGCCAGGCGGCACATCAAGGCCAAGCCCG
CCGATGTGCCCGCACCGCAGGCCAAGGCTGCGGAACCCGCGCCGGCACCCAAGACGCCGGAGCC
ACGGCGGCCGAAGCAGGGGGGCAAGGCTGAAAAGCCGGCCCCCGCTGCGGCCCCGACCGGCTTC
ACCTTCAACCCAACACCGGACAAAAAGGATCCAGCAGTTCAACCTGTTGATAGTACGTACTAAG
CTCTCATGTTTCACGTACTAAGCTCTCATGTTTAACGTACTAAGCTCTCATGTTTAACGAACTA
AACCCTCATGGCTAACGTACTAAGCTCTCATGGCTAACGTACTAAGCTCTCATGTTTCACGTAC
TAAGCTCTCATGTTTGAACAATAAAATTAATATAAATCAGCAACTTAAATAGCCTCTAAGGTTT
TAAGTTTTATAAGAAAAAAAAGAATATATAAGGCTTTTAAAGCTTTTAAGGTTTAACGGTTGTG
GACAACAAGCCAGGGATGTAACGCACTGAGAAGCCCTTAGAGCCTCTCAAAGCAATTTTCAGTG
ACACAGGAACACTTAACGGCTGACATGGCCTTAGAAGGCGTTTTAGAAGCCCTAATGGATGGAG
TCAAAACAAACTAGCGATCGAATTCCCGGGAGAGCTC

Appendix 2: pGOA1193 sequence

pIVET5n (pGOA1193): 6,929 bp. tnpR 39-590, lacZ 605-3679, traJ 4669-4298, traK* (extra amino acids fused) 5037-5528, oriR6K 5434-5917, bla 6032-6892.

CTCGAGATCTCAATTGTTCGAATTTAGGATACATTTTTATGCGACTTTTTGGTTACGCACGGGT
ATCAACCAGCCAGCAATCTCTCGATATTCAGGTTCGGGCACTCAAAGACGCAGGCGTGAAAGCA
AATCGCATCTTTACTGACAAGGCATCGGGCAGTTCATGCGATCGGAAAGGGCTGGACTTGCTGA
GGATGAAGGTGGAGGAAGGTGACGTCATCTTGGTGAAGAAACTTGACCGCCTTGGGCGCGATAC
TGCTGACATGATCCAGTTAATAAAAGAGTTTGACGCCCAAGGTGTATCCATTCGGTTTATTGAT
GACGGAATCAGTACCGATGGGGAGATGGGTAAAATGGTTGTCACTATTCTATCTGCAGTGGCCC
AGGCAGAACGACAGAGAATACTAGAGCGTACCAATGAAGGTCGCCAAGAGGCAATGGCAAAAGG
AGTTGTTTTTGGTAGAAAAAGAAAAATAGATAGAGATGCAGTATTAAATATGTGGCAACAGGGG
TTAGGTGCCTCACATATATCAAAAACAATGAATATTGCTCGTTCAACAGTATATAAAGTAATAA
ATGAAAGCAACTAAGGAGGAAAGTCGACATGACCATGATTACGGATTCACTGGCCGTCGTTTTA
CAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTT
TCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCT
GAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGCGGTGCCGGAAAGCTGGCTGGAG
TGCGATCTTCCTGAGGCCGATACTGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTTACGATG
CGCCCATCTACACCAACGTAACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAA
TCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACG
CGAATTATTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAACGGGCGCTGGGTCGGTT
ACGGCCAGGACAGTCGTTTGCCGTCTGAATTTGACCTGAGCGCATTTTTACGCGCCGGAGAAAA
CCGCCTCGCGGTGATGGTGCTGCGTTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGG
CGGATGAGCGGCATTTTCCGTGACGTCTCGTTGCTGCATAAACCGACTACACAAATCAGCGATT
TCCATGTTGCCACTCGCTTTAATGATGATTTCAGCCGCGCTGTACTGGAGGCTGAAGTTCAGAT
GTGCGGCGAGTTGCGTGACTACCTACGGGTAACAGTTTCTTTATGGCAGGGTGAAACGCAGGTC
GCCAGCGGCACCGCGCCTTTCGGCGGTGAAATTATCGATGAGCGTGGTGGTTATGCCGATCGCG
TCACACTACGTCTGAACGTCGAAAACCCGAAACTGTGGAGCGCCGAAATCCCGAATCTCTATCG
TGCGGTGGTTGAACTGCACACCGCCGACGGCACGCTGATTGAAGCAGAAGCCTGCGATGTCGGT
TTCCGCGAGGTGCGGATTGAAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATTCGAG
GCGTTAACCGTCACGAGCATCATCCTCTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGCA
GGATATCCTGCTGATGAAGCAGAACAACTTTAACGCCGTGCGCTGTTCGCATTATCCGAACCAT
CCGCTGTGGTACACGCTGTGCGACCGCTACGGCCTGTATGTGGTGGATGAAGCCAATATTGAAA
CCCACGGCATGGTGCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCGATGAGCGA
ACGCGTAACGCGAATGGTGCAGCGCGATCGTAATCACCCGAGTGTGATCATCTGGTCGCTGGGG
AATGAATCAGGCCACGGCGCTAATCACGACGCGCTGTATCGCTGGATCAAATCTGTCGATCCTT
CCCGCCCGGTGCAGTATGAAGGCGGCGGAGCCGACACCACGGCCACCGATATTATTTGCCCGAT
GTACGCGCGCGTGGATGAAGACCAGCCCTTCCCGGCTGTGCCGAAATGGTCCATCAAAAAATGG
CTTTCGCTACCTGGAGAGACGCGCCCGCTGATCCTTTGCGAATACGCCCACGCGATGGGTAACA
GTCTTGGCGGTTTCGCTAAATACTGGCAGGCGTTTCGTCAGTATCCCCGTTTACAGGGCGGCTT
CGTCTGGGACTGGGTGGATCAGTCGCTGATTAAATATGATGAAAACGGCAACCCGTGGTCGGCT
TACGGCGGTGATTTTGGCGATACGCCGAACGATCGCCAGTTCTGTATGAACGGTCTGGTCTTTG
CCGACCGCACGCCGCATCCAGCGCTGACGGAAGCAAAACACCAGCAGCAGTTTTTCCAGTTCCG
TTTATCCGGGCAAACCATCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGATAACGAGCTC
CTGCACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGTCG
CTCCACAAGGTAAACAGTTGATTGAACTGCCTGAACTACCGCAGCCGGAGAGCGCCGGGCAACT

CTGGCTCACAGTACGCGTAGTGCAACCGAACGCGACCGCATGGTCAGAAGCCGGGCACATCAGC
GCCTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCCGCCGCGTCCCACGCCA
TCCCGCATCTGACCACCAGCGAAATGGATTTTTGCATCGAGCTGGGTAATAAGCGTTGGCAATT
TAACCGCCAGTCAGGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACAACTGCTGACGCCG
CTGCGCGATCAGTTCACCCGTGCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCCGCA
TTGACCCTAACGCCTGGGTCGAACGCTGGAAGGCGGCGGGCCATTACCAGGCCGAAGCAGCGTT
GTTGCAGTGCACGGCAGATACACTTGCTGATGCGGTGCTGATTACGACCGCTCACGCGTGGCAG
CATCAGGGGAAAACCTTATTTATCAGCCGGAAAACCTACCGGATTGATGGTAGTGGTCAAATGG
CGATTACCGTTGATGTTGAAGTGGCGAGCGATACACCGCATCCGGCGCGGATTGGCCTGAACTG
CCAGCTGGCGCAGGTAGCAGAGCGGGTAAACTGGCTCGGATTAGGGCCGCAAGAAAACTATCCC
GACCGCCTTACTGCCGCCTGTTTTGACCGCTGGGATCTGCCATTGTCAGACATGTATACCCCGT
ACGTCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGACGCGCGAATTGAATTATGGCCCACACCA
GTGGCGCGGCGACTTCCAGTTCAACATCAGCCGCTACAGTCAACAGCAACTGATGGAAACCAGC
CATCGCCATCTGCTGCACGCGGAAGAAGGCACATGGCTGAATATCGACGGTTTCCATATGGGGA
TTGGTGGCGACGACTCCTGGAGCCCGTCAGTATCGGCGGAATTCCAGCTGAGCGCCGGTCGCTA
CCATTACCAGTTGGTCTGGTGTCAAAAATAATAAATCCCGCGGTGGCGGCCGCTCTAGCCCCAG
CCGACCAGGCTTTCCACGCCCGCGTGCCGCTCCATGTCGTTCGCGCGGTTCTCGGAAACGCGCT
GCCGCGTTTCGTGATTGTCACGCTCAAGCCCGTAGTCCCGTTCGAGCGTCGCGCAGAGGTCAGC
GAGGGCGCGGTAGGCCCGATACGGCTCATGGATGGTGTTTCGGGTCGGGTGAATCTTGTTGATG
GCGATATGGATGTGCAGGTTGTCGGTGTCGTGATGCACGGCACTGACGCGCTGATGCTCGGCGA
AGCCAAGCCCAGCGCAGATGCGGTCCTCAATCGCGCGCAACGTCTCCGCGTCGGGCTTCTCTCC
CGCGCGGAAGCTAACCAGCAGGTGATAGGTCTTGTCGGCCTCGGAACGGGTGTTGCCGTGCTGG
GTCGCCATCACCTCGGCCATGACAGCGGGCAGGGTGTTTGCCTCGCAGTTCGTGACGCGCACGT
GACCCAGGCGCTCGGTCTTGCCTTGCTCGTCGGTGATGTACTTCACCAGCTCCGCGAAGTCTCT
CTTCTTGATGGAGCGCATGGGGACGTGCTTGGCAATCACGCGCACCCCCCGGCCGTTTTAGCGG
CTAAAAAAGTCATGGCTCTGCCCTCGGGCGGACCACGCCCATCATGACCTTGCCAAGCTCGTCC
TGCTTCTCTTCGATCTTCGCCAGCAGGGCGAGGATCGTGGCATCACCGAACCGCGCCGTGCGCG
GGTCGTCGGTGAGCCAGAGTTTCAGCAGGCCGCCCAGGCGGCCCAGGTCGCCATTGATGCGGGC
CAGCTCGCGGACGTGCTCATAGTCCACGACGCCCGTGATTTTGTAGCCCTGGCCGACGGCCAGC
AGGTAGGCCGACAGGCTCATGCCGGCCGCCGCCGCCTTTTCCTCAATCGCTCTTCGTTCGTCTG
GAAGGCAGTACACCTTGATAGGTGGGCTGCCCTTCCTGGTTGGCTTGGTTTCATCAGCCATCCG
CTTGCCCTCATCTGTTACGCCGGCGGTAGCCGGCCAGCCTCGCAGAGCAGGATTCCCGTTGAGC
ACCGCCAGGTGCGAATAAGGGACAGTGAAGAAGGAACACCCGCTCGCGGGTGGGCCTACTTCAC
CTATCCTGCCCGGCTGACGCCGTTGGATACACCAAGGAAAGTCTACACGAACCCTTTGGCAAAA
TCCTGTATATCGTGCGAAAAAGGATGGATATACCGAAAAAATCGCTATAATGACCCCGAAGCAG
GGTTATGCAGCGGAAAAGCGCTGCTTCCCTGCTGTTTTGTGGAATATCTACCGACTGGAAACAG
GCAAATGCAGGAAATTACTGAACTGAGGGGACAGGCGAGAGACGATGCCAAAGAGCTACACCGA
CGAGCTGGCCGAGTGGGTTGAATCCCGCGCGGCCAAGAAGCGCCGGCGTGATGAGGCTGCGGTT
GCGTTCCTGGCGGTGGGGGCGGATGTCGAGGCGGCGTTAGCGTCCGGCTATGCGCTCGTCACCA
TTTGGGAGCACATGCGGGAAACGGGGAAGGTCAAGTTCTCCTACGAGACGTTCCGCTCGCACGC
CAGGCGGCACATCAAGGCCAAGCCCGCCGATGTGCCCGCACCGCAGGCCAAGGCTGCGGAACCC
GCGCCGGCACCCAAGACGCCGGAGCCACGGCGGCCGAAGCAGGGGGGCAAGGCTGAAAAGCCGG
CCCCCGCTGCGGCCCCGACCGGCTTCACCTTCAACCCAACACCGGACAAAAAGGGGCGGATCCG
GCCACGATGCGTCCGGCGTAGAGGATCTGAAGATCAGCAGTTCAACCTGTTGATAGTACGTACT
AAGCTCTCATGTTTCACGTACTAAGCTCTCATGTTTAACGTACTAAGCTCTCATGGCTAACGTA
CTAAGCTCTCATGTTTCACGTACTAAGCTCTCATGTTTGAACAATAAAATTAATATAAATCAGC
AACTTAAATAGCCTCTAAGGTTTTAAGTTTTATAAGAAAAAAAAGAATATATAAGGCTTTTAAA

GCTTTTAAGGTTTAACGGTTGTGGACAACAAGCCAGGGATGTAACGCACTGAGAAGCCCTTAGA
GCCTCTCAAAGCAATTTTGAGTGACACAGGAACACTTAACGGCTGACATGGGAATTCCACATGT
GGAATTCCACATGTGGAATTGTGAGCGGATAACAATTTGTGGAATTCCCGGGAGAGCTCGATAT
CGCATGCGGTACCTCTAGAAGAAGCTTGGGATCCGCCCCGCGGAACCCCTATTTGTTTATTTTT
CTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATAT
TGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCAT
TTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTT
GGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGC
CCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCC
GTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGA
GTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCT
GCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGG
AGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGA
GCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACG
TTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGA
TGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGC
TGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGT
AAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATA
GACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAATCGTACGAGCCGGTGCTAT
CAAAGCACCGGCTTTTT